

## Development of Somatic Complaints Among Adolescents and Young Adults in Switzerland

André Berchtold\*, Joan-Carles Surís\*\*, Thomas Meyer\*\*\*,  
and Zhivko Taushanov\*

**Abstract:** In this study we explored the development of somatic complaints among adolescents and young adults aged 16 to 30 years in Switzerland. Using data from the Transitions from Education to Employment (TREE) study, we applied a hidden Markovian model with covariates to cluster trajectories representing the sum of eight somatic complaints. The resulting groups differed mainly in terms of gender, reading literacy, and substance use. The trajectories of somatic complaints were also related to the number of critical events experienced by the respondents.

**Keywords:** trajectories, somatic complaints, critical life events, clustering, adolescence

### Entwicklung von somatischen Beschwerden bei Jugendlichen und jungen Erwachsenen in der Schweiz

**Zusammenfassung:** Der Beitrag untersucht, wie sich in der Schweiz somatische Beschwerden bei Jugendlichen bzw. jungen Erwachsenen im Alter zwischen 16 und 30 Jahren entwickeln. Mit den Daten der TREE-Studie wurde ein verstecktes Markov-Modell mit Kovariaten gerechnet, um Verläufe von acht somatischen Beschwerden zu clustern. Die so gebildeten Gruppen unterscheiden sich signifikant bezüglich Geschlecht, kognitiven Fertigkeiten und Substanzkonsum. In den untersuchten Verläufen zeigen sich auch Zusammenhänge mit kritischen Lebensereignissen der Befragten.

**Schlüsselwörter:** Verläufe, somatische Beschwerden, kritische Lebensereignisse, Clustering, Jugend

### Développement de plaintes somatiques chez les adolescents et jeunes adultes en Suisse

**Résumé:** Nous avons exploré le développement de plaintes somatiques chez les adolescents et jeunes adultes âgés de 16 à 30 ans vivant en Suisse. Sur la base de données de l'enquête TREE (Transitions from Education to Employment), nous avons appliqué un modèle markovien caché avec covariables afin de classifier des trajectoires représentant la somme de huit plaintes somatiques. Les groupes obtenus diffèrent en termes de genre, de niveau de lecture et de consommation de substances. Les trajectoires de plaintes somatiques sont aussi liées au nombre d'événements de vie critiques dont les personnes interrogées ont fait l'expérience.

**Mots-clés:** trajectoires, plaintes somatiques, événements de vie critiques, classification, adolescence

\* Institute of Social Sciences & National Centre of Competence in Research LIVES, University of Lausanne, CH-1015 Lausanne, andre.berchtold@unil.ch, zhivko.taushanov@unil.ch.

\*\* Institute of Social and Preventive Medicine and General Pediatrics, Lausanne University Hospital, CH-1010 Lausanne, joan-carles.suris@chuv.ch.

\*\*\* Institute of Sociology, University of Bern, CH-3012 Bern, thomas.meyer@soz.unibe.ch.

## 1 Introduction<sup>1</sup>

Somatic complaints such as headaches, stomach aches or sleep disturbances are very common at all ages. They are a leading reason for seeking medical care, accounting for up to 50% of new medical outpatient visits (Mohapatra et al. 2014). These symptoms often appear during childhood and then increase through adolescence and adulthood. In Switzerland, a study showed an increase in the number and importance of these symptoms among 11–15 year olds between 1996 and 2004 (Dey et al. 2015). In addition to lowering the quality of everyday life, the presence of somatic complaints is often a clue for more serious problems, either already present or likely to grow rapidly. Understanding the causes of somatic complaints is therefore crucial to prevent and/or identify and treat more important health problems. Accordingly, a Dutch study showed that young adults with severe somatic disabilities since childhood achieved less life milestones than their healthy peers, or achieved them with delay, implying a lower probability of full social and professional integration (Verhoof et al. 2012). A higher level of somatic issues at mid-age has also been associated with reduced accumulation of social capital from adolescence throughout the life course (Jonsson et al. 2014).

Several studies have established a relationship during adolescence between somatic complaints and negative affectivity (Vassend 1989), depression or other mental illnesses (Härmä et al. 2002; Kapfhammer 2006). Even though somatic symptoms are known to often predict mental problems or diseases during the life course (Nakao and Yano 2006; Bekhuis et al. 2016), there is no robust evidence to date as to whether depression is a cause or a consequence of somatic complaints, or if the two phenomena appear coincidentally (Goodwin 2006).

While specific causes can sometimes be found, in many situations the presence of somatic complaints remains unexplained, and their appearance and disappearance cannot be related to clear medical causes. Exogenous and environmental causes therefore need to be explored. In their review paper, Barsky et al. (2001) showed that, already during adolescence, women tend to report a larger number of somatic symptoms than men, both in medical and general populations; this difference being likely to be already present during childhood. However, similar to what is observed for depression (Sigmon et al. 2005), this difference could also be explained by a gender-related reporting bias, with women having a lower reporting threshold than men. Children having to cope with social and family difficulties (e.g. parental substance abuse) also reported more somatic complaints (Radke-Yarrow and Brown 1993), as well as children living in a generally less than average socio-

---

1 This publication benefited from the support of the Swiss National Centre of Competence in Research LIVES – Overcoming vulnerability: Life course perspectives, which is financed by the Swiss National Science Foundation (grant number: 51NF40-160590). The authors are grateful to the Swiss National Science Foundation for its financial assistance.

economic environment, the impact of the latter being recognized well beyond the medical literature (Choi 2009).

Critical life events such as the death of a parent, the birth of a child or an accident are also likely to have an impact on somatic health. In a study among Finnish high school students, somatic symptoms were positively correlated with critical life events (Poikolainen et al. 1995). Similarly, school or work difficulties and stress are considered as possible explanations for the early development of somatic complaints (Sek-yum et al. 1992; Hart et al. 2013). When life transitions or critical events occur, more support is required, because these periods can be phases of increased vulnerability. This process was put into evidence by Compas et al. (1986) in their longitudinal study among high school and college students. These findings are linked to the concept of vulnerability that, following the work of Spini et al. (2017), can be defined as a weakening process coupled with a lack of resources in one or several life domains, such as education, family support or wealth. Somatic complaints could then be considered as an additional warning message mediating the relationship between critical life events and vulnerability.

In addition to life events over which we have little control (death of a relative, accident, etc.), risk taking behaviour such as substance abuse or internet overuse also need to be considered (Surís et al. 2014). Alcohol and illegal drug use have been associated with somatic complaints for both genders, while smoking has been associated with somatic complaints for males only (Poikolainen et al. 1995). The relationship between substance use disorders and somatic complaints is complex and may differ depending on the social or cultural background of the population under study (Yoshimasu 2012).

Even though a large body of literature has already investigated somatic complaints, most of these publications are based on cross-sectional data taken for instance from HBSC<sup>2</sup> or ESPAD<sup>3</sup> databases. Moreover, a substantive amount of past studies considered specific populations, mainly those mentally ill or susceptible to developing some kind of mental illness. However, somatic complaints concern the whole population. For instance, in a recent Danish study covering the general adult population, it was found that 94.9% of the respondents presented at least one somatic symptom (Eliassen et al. 2016). In a Swiss study among 8<sup>th</sup> graders, 14.2% of respondents reported back problems and 16.7% reported headaches (Surís et al. 2014). A few longitudinal studies can still be cited: Wright and Wright (1981) studied the frequency of somatic complaints among 90 adolescents observed four times between 11 and 18 years. Barkmann et al. (2015) used a much larger sample of 2857 children and adolescents aged between 7 and 17 at baseline, but they followed them for only three years. Thus, there is still a lack of longitudinal studies using a large sample of the general population of adolescents and following them

2 Health Behaviour in School-Aged Children. Website: [www.hbsc.org](http://www.hbsc.org).

3 The European School Survey Project on Alcohol and Other Drugs. Website: [www.espad.org](http://www.espad.org).

individually for a long time. Our study then contributes to broadening the field of research concerning the development of somatic complaints, thereby filling a substantial gap in the literature.

The main purpose of this study was to assess the presence and development of somatic complaints among adolescents and young adults living in Switzerland during a life period during which crucial transitions occur, such as entry on the labour market and the foundation of a family. With regard to the development of individuals from mid-adolescence to young adulthood, we searched to identify specific subgroup trajectories of somatic complaint development and link these trajectories both to personal and socio-economic factors prone to shape the overall trajectory, as well as to critical life events. Given the scarcity of previous longitudinal studies analysing somatic complaints among adolescents and young adults, it was difficult to predict which would be the most likely shapes of these trajectories, except maybe, following Barkmann et al. (2015), a slight trend to an increase of the number of complaints during early adolescence associated with a high variability between individuals. Moreover, it was reasonable to postulate that some adolescents experience only a very small number of complaints, which corresponds to a quite flat trajectory with a low average value. We postulated then that not everybody is equal regarding the overall level of somatic complaints and their development over time. More specifically, we hypothesized that 1) females, 2) living in a low SES context, and 3) lower academic achievement were all associated with an increased prevalence of somatic symptoms; and that 4) critical life events were related to sudden changes in the level of reported somatic symptoms. We further hypothesized that 5) risky behaviours such as substance use were also associated with specific trajectories of somatic complaint development, but without being a trigger for the onset or change of these symptoms.

## 2 Methods

We describe in this section the dataset used in our study and the different statistical analyses performed first to classify the trajectories of somatic complaints, and then to interpret the results.

### 2.1 Data

We used data from the Transitions from Education to Employment (TREE) study (TREE 2016)<sup>4</sup>. TREE is a follow-up survey of the Swiss sample tested by the Programme for International Student Assessment (PISA) survey in 2000, collecting longitudinal data among more than 6000 school leavers from 2001 (mean

<sup>4</sup> The Swiss panel study TREE is a social science data infrastructure mainly funded by the Swiss National Science Foundation (SNF) and located at the University of Bern.

age 16 years) to 2014. Data available to date include PISA 2000 (baseline survey) and nine follow-up panel waves carried out between 2001 and 2014 (at annual intervals between 2001 and 2007, then in 2010 and 2014), but the study is still ongoing and a further wave is planned for 2019. The TREE study was not submitted to an ethics committee, since it was not required by the Swiss law on human research. However, all participants were required to give their express consent. The presence of eight somatic complaints (stomach ache, lack of appetite, lack of concentration, vertigo, sleeping disorder, nervousness, fatigue, headaches) was regularly surveyed at each TREE panel wave, drawing on the *Berner Fragebogen zum Wohlbefinden Jugendlicher* (Grob et al. 1991). There were five possible answers for each somatic complaint ranging from never to every day. These answers were recoded from 0 to 4 and then summed in order to obtain an overall somatic complaint score ranging from 0 to 32. This sum score was then used as the dependent variable in our analyses. Higher scores can indicate either a larger number of reported complaints or a higher intensity of some complaints, but most scores above 10 indicate that at least one complaint is at an important level.

Several covariates, operationalizing the key determinants for somatic complaints identified in the literature, were included in our model, either fixed or time-varying. The fixed covariates were gender (female/male), country of birth (Switzerland/other), academic track attended at mandatory school (high/extended/basic), residence (rural area/urban area), PISA reading literacy (6 levels from 0 = very low to 5 = very high; treated as a numerical scale hereafter), highest parental socio-economic status and family wealth. Residence was included based on the hypothesis that living in an urban area can be more stressful than living in a rural one, which could in turn favour the development of somatic complaints. Literacy can be considered as an indicator of the psychological state of an individual, and it can also be related to the socio-economic status. To take into account the possible influence of a low socio-economic environment, we considered two different ways of measuring the standard of living: socio-economic status and family wealth. Socio-economic status was operationalized using the highest parental International Socio-Economic Index (ISEI) introduced by Ganzeboom et al. (1992), a scale ranging from 10 to 90, higher values indicating higher socio-economic status. Family wealth was measured through a scale representing the possessions of the family such as cars and TV sets. The scale was normalized with the zero value indicating an average wealth. All these variables were measured in 2000 as part of the PISA survey. In contrast, critical life events were measured at each subsequent wave of the TREE survey. The number of surveyed critical life events varied between 12 and 16 across panel waves, including an open text option from the second wave onward. Reported events comprised relocation of parental family; moving out of the parental home; parental and own separation or divorce; death, serious accident/illness or unemployment of relevant others; trouble with the police; unhappy relationship; serious conflicts in the family, at school or

at work; pregnancy and parenthood. Two time-varying covariates were computed from these life events. The first one was the number of critical life events reported each year. The second one was a dichotomous factor indicating whether at least one critical event was reported or not. Following the existing literature, we postulated that each covariate could have a direct impact on somatic complaints. However, we performed a distinction between covariates observed only once, in 2000, and covariates observed continuously from 2001 to 2014. The latter ones were introduced in the statistical model to directly influence the average level of somatic complaints of an individual, but the previous ones were used at the latent level, giving them a more persistent effect throughout the period of observation. Finally, the consumption level of four types of substances (alcohol, tobacco, cannabis, tranquilizers/sleeping pills) was assessed at each TREE wave with five possible answers ranging from “never” to “every day,” but as explained in the next Section, these variables were not used directly into the statistical model.

## 2.2 Statistical analysis

For the statistical analysis, we relied on the Hidden Mixture Transition Distribution Model (HMTD), a complete Markovian framework for the clustering, modelling and analysis of sequences of continuous data. This two-level model builds on the Double Chain Markov Model, another Markovian model previously developed for categorical variables (Berchtold 2002). The HMTD model combines a hidden and an observed level. At the hidden level, a latent variable  $X$  taking values in a finite set  $(1, \dots, k)$  was used to represent different possible kinds of evolution of somatic complaints (or states in the HMTD terminology). To the best of our knowledge, the only competitor to the HMTD model given the requirements of our study was the Growth Mixture Model (GMM), but we preferred to rely on HMTD because of its flexibility. In the standard HMTD model, similar to a hidden Markov model, a finite transition matrix was used to represent all possible transitions between the  $k$  states. However, since we used the model as a clustering tool, the transition matrix was restricted to the identity one, implying that each sequence belonged to the same state from the beginning to the end. The model then produced a clustering of the data sequences into  $k$  mutually exclusive groups. Moreover, the model allowed for a set of time-invariant covariates to influence the probability of any sequence to be assigned to a specific group. In practice, a multinomial regression using the covariates as explanatory variables, and the current clustering of the observed sequences as the dependent variable, was computed during each reestimation of the model parameters, and the resulting probabilities were then used as initial probabilities of belonging to each group. In function of the combination of covariates, each observed sequence of somatic complaints could have a different set of initial probabilities.

At the visible level, a Gaussian distribution was assigned to each hidden state of the model. The mean and the variance of each distribution were modelled through

an autoregressive model, with the possible inclusion of covariates. We chose to model only the mean of each distribution, the variance being constant. Moreover, we added a time-varying covariate representing the occurrence of critical life events. For instance, the following equation represented the modelling of the mean level of somatic complaints at time  $t$ , using two lags of the somatic complaints dependent variable  $y$  and a covariate  $c$ :

$$\mu_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 c_t$$

To summarize, at the visible level the model predicts the current level of somatic complaints in function of the last two observations of this same variable, and of the number of critical life events experienced during the same period. The hidden model is used to assign each respondent to one specific group in function of a set of fixed covariates. For more details, the HMTD was completely described in Bolano and Berchtold (2016), and its estimation was discussed in Taushanov and Berchtold (2017). At the visible level, confidence intervals are obtained using a bootstrap procedure.

In the first step, the HMTD model was used to identify the required number of groups for classifying the data sequences, and the order of dependence for the autoregressive modelling of the mean value of the somatic complaints scale. No covariates were used at this point. Models were compared on the basis of their log-likelihood, their Bayesian Information Criterion (BIC) values (Raftery 1995), and the number of sequences assigned to each group, thereby discarding solutions with very few sequences in some groups. Then, time-invariant covariates were introduced one by one at the hidden level. All significant covariates were then introduced simultaneously in the model to improve the clustering. Finally, the time-varying covariate representing the occurrence of critical life events was added at the visible level to improve the modelling of the mean of the somatic complaint sum score. No averaging was performed across periods for the life event variables, so an individual could have a different value on each observation occasion.

Critical life events were mainly shocks occurring at a precise time, but whose effects could be felt for a long period. Specific examples were the death of a family member, an unhappy love or a sudden hospitalization. Their impact on the sum of somatic complaints could therefore be easily conceptualized. On the other hand, substance use was mostly a continuous behaviour that was difficult to break into specific events. Even the beginning of consumption of a specific substance was difficult to assess, because 1) someone could begin (and cease) to use a substance several times, and 2) our dataset could not be used to determine whether a specific substance was used before the first wave. Moreover, a sudden change in the level of consumption could not always be clearly identified in our data, because questions about substance use only referred to the month preceding the survey panel. Therefore, we chose to integrate critical life events and substance use in two different ways

in our analyses. Critical events were used as a time-varying covariate influencing the average level of somatic complaints into each group of the clustering, when the association between substance use and trajectories was established a posteriori on the final clustering using a chi-square test.

Continuous covariates were standardized in order to ease the convergence of the optimization algorithm. The Type I error was fixed to 5%. All computations were performed using the R open source statistical software, especially the ad hoc estimation algorithm described in Taushanov and Berchtold (2017).

3 Results

Data from N = 1 161 respondents continuously observed from 2001 to 2014 were included in all analyses. These individuals represent only 18% of the total TREE sample, but as it will be discussed later on, we preferred not to impute missing data, and this choice had only little influence on our results. Table 1 summarizes the main information about the sum score for somatic complaints. This scale showed good psychometric properties with a Cronbach’s alpha value ranging from 0.78 (T1) to 0.82 (T6). Whatever the wave, the score was highly variable from one respondent to another, but the central tendency measured by the mean and the median did not vary much. Most scores were below 20, but each year a small number of larger values were observed.

Table 1 Main characteristics of the somatic complaint score

Coefficients	Year								
	2001	2002	2003	2004	2005	2006	2007	2010	2014
Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Maximum	32.00	29.00	28.00	28.00	31.00	32.00	27.00	25.00	28.00
Median	6.00	6.00	6.00	6.00	5.00	6.00	6.00	6.00	5.00
Mean	7.14	6.91	7.07	6.75	6.27	7.14	6.85	6.26	6.05
Standard deviation	4.83	4.81	4.79	4.76	4.58	4.76	4.59	4.31	4.24
Cronbach’s alpha	0.78	0.80	0.80	0.80	0.80	0.82	0.80	0.79	0.79

Table 2 and Table 3 describe the covariates considered in this study. When comparing our data with the full TREE sample, females and students in the pre-gymnasial school track were over-represented, while German speaking youths were slightly under-represented. This will be discussed later, but it did not affect our results. The number of critical life events increased year to year, which is to be expected



given the probability of experiencing some of the events surveyed, such as getting married or becoming a parent, increases with age. Moreover, the variability between respondents also increased with age, even though few individuals reported many events in a given survey wave.

Table 2 Main characteristics of the time invariant covariates measured in year 2000

Variable	Categories	Distribution
Gender	Female	749 (64.51%)
	Male	412 (35.49%)
Country of birth	Switzerland	1097 (94.49%)
	Other	64 (5.51%)
Academic track attended at lower secondary education level	High	574 (49.44%)
	Extended	432 (37.21%)
	Basic	155 (13.35%)
Residence	Rural area	433 (37.30%)
	Urban area	728 (62.70%)
(PISA) Reading literacy	Min.: 0, Max.: 5	3.50 (1.00)
Highest parental ISEI	Min.: 16, Max.: 90	53.18 (15.45)
Family wealth	Min.: -2.93, Max.: -3.38	0.05 (0.76)

Note: We provide the prevalence of each category and the corresponding percentage for categorical variables, and the mean and standard deviation for numerical variables.

Table 3 Main characteristics of the critical life events score

Coefficients	Year								
	2001	2002	2003	2004	2005	2006	2007	2010	2014
Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Maximum	6.00	5.00	7.00	12.00	6.00	8.00	8.00	14.00	9.00
Median	0.00	0.00	0.00	0.00	1.00	1.00	1.00	2.00	2.00
Mean	0.69	0.69	0.70	0.74	0.91	1.09	1.09	2.26	2.57
Standard deviation	0.97	0.93	0.94	1.06	1.11	1.26	1.23	1.50	1.50
Number of respondents reporting > 0 events	518	526	536	547	605	664	680	1041	1085

As a first step, we considered HMTD models with 2 to 8 hidden groups and a first- or second-order dependence for the mean of the somatic complaint score. Based on the BIC, the preferred model was the second-order model with 6 groups. Using the previous two observations of the dependent variable to explain the current observations yielded better results than using only the immediately preceding observation. We subsequently added the fixed covariates one by one at the hidden

level. Five covariates contributed to improve the fit of the model: gender, residence, reading literacy, socio-economic status and family wealth. These covariates were then introduced together at the hidden level, and we added the time-varying critical life events covariate at the visible level, either in its continuous or dichotomous form. Both versions of this latter covariate proved useful in improving the clustering of the somatic complaint score trajectories, but the best results were obtained with the continuous covariate. Finally, since two of the six groups were very close in terms of trajectories and parameters, we computed the same model with only five groups. This model was chosen as the final solution.

Table 4 displays the parameters of the final HMTD model, and Figure 1 shows the clustering of the somatic complaint trajectories into the five groups identified by the model. Trajectories must be analysed in terms of average value and of vari-

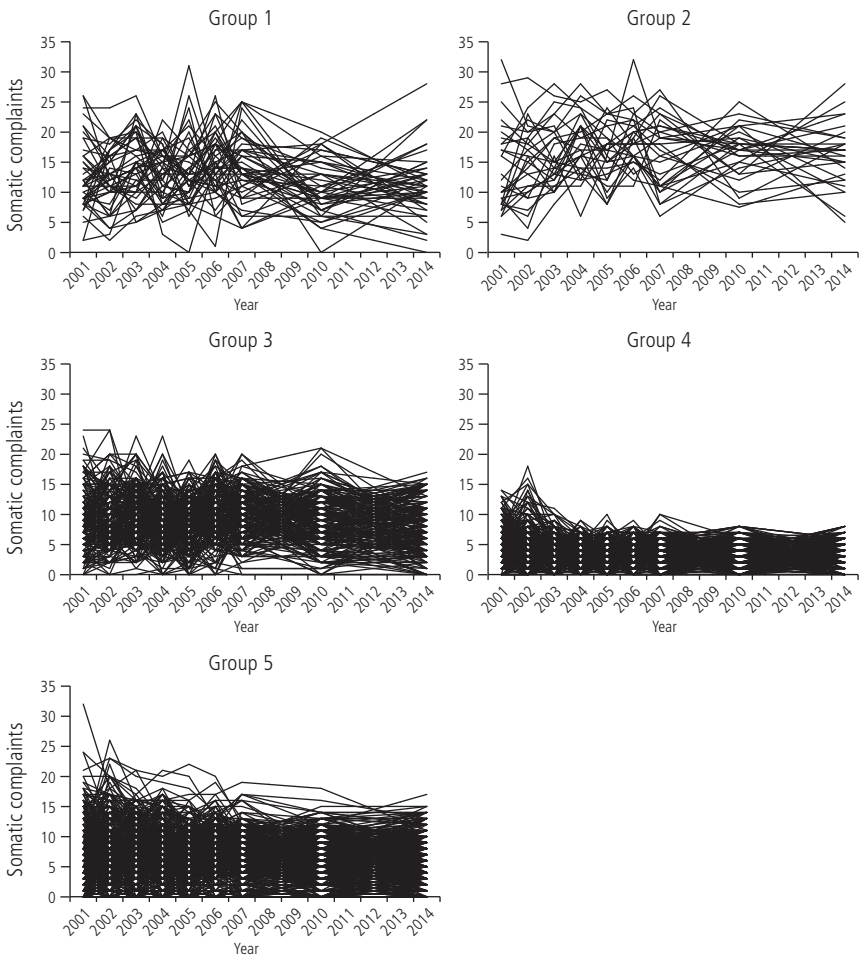
Table 4                      Parameters of the final HMTD clustering model

Hidden level: Clustering of somatic complaints trajectories into 5 groups					
Groups	Residence (urban area)	Gender (male)	Reading literacy	Hisei	Family wealth
1	0.42 [−0.08; 0.93]	<b>−1.00</b> [−1.59; −0.41]	−0.13 [−0.39; 0.12]	0.06 [−0.18; 0.31]	0.16 [−0.16; 0.47]
2	0.43 [−0.25; 1.12]	<b>−2.10</b> [−3.29; −0.91]	<b>−0.47</b> [−0.81; −0.12]	−0.03 [−0.36; 0.30]	0.13 [−0.29; 0.56]
3	0.02 [−0.32; 0.35]	−0.30 [−0.65; 0.05]	−0.14 [−0.32; 0.04]	0.03 [−0.14; 0.21]	0.04 [−0.19; 0.26]
4	0.03 [−0.26; 0.31]	<b>0.44</b> [0.15; 0.72]	−0.02 [−0.17; 0.13]	0.03 [−0.12; 0.18]	0.18 [−0.01; 0.37]
5	Ref.	Ref.	Ref.	Ref.	Ref.

Visible level: Observed levels of somatic complaints					
Groups	Variance	Factors explaining the mean level of somatic complaints			
		Constant	Lag 1	Lag 2	Critical life events
1	<b>26.99</b>	<b>11.68</b>	<b>0.20</b>	−0.09	−0.08
(n = 46)	[21.31; 32.56]	[10.15; 13.62]	[0.09; 0.33]	[−0.22; 0.01]	[−0.47; 0.39]
2	<b>20.55</b>	<b>9.55</b>	<b>0.16</b>	<b>0.29</b>	0.17
(n = 30)	[15.48; 25.34]	[7.90; 11.99]	[0.04; 0.26]	[0.15; 0.41]	[−0.39; 0.68]
3	<b>18.01</b>	<b>6.29</b>	<b>0.28</b>	0.01	0.14
(n = 204)	[16.92; 19.06]	[5.79; 6.92]	[0.24; 0.33]	[−0.05; 0.06]	[−0.10; 0.36]
4	<b>3.39</b>	<b>2.25</b>	<b>0.23</b>	<b>0.11</b>	<b>0.14</b>
(n = 353)	[3.18; 3.57]	[2.06; 2.45]	[0.20; 0.27]	[0.08; 0.15]	[0.04; 0.23]
5	<b>7.21</b>	<b>1.71</b>	<b>0.42</b>	<b>0.31</b>	0.05
(n = 528)	[6.80; 7.62]	[1.48; 1.96]	[0.38; 0.46]	[0.27; 0.34]	[−0.05; 0.16]

Note: At the hidden level, the last group served as reference for the computation of the multinomial regression used to add the fixed covariates to the model. We provide for each parameter the point estimation and the 95% confidence interval. Parameters significant at the 95% level are printed in bold.

Figure 1 Somatic complaint trajectories of the final five groups identified by the model



ability, both during a particular sequence or between sequences. Accordingly, the figure and the model parameters indicated that 1) the groups differed both in terms of average level of somatic complaints and variability; 2) inter-subject variability remained high, even within the same cluster, indicating that almost all individuals followed their own trajectory; and 3) intra-subject variability (that is across time for a specific individual) was high for the trajectories classified in groups 1 to 3, and much lower for trajectories classified in groups 4 and 5. Group 5, which comprises

about half the sample ( $n = 528$ ), was used as the reference group for the analysis. Regarding the covariates used at the hidden level, gender was the most important one for distinguishing between the groups, with more males in group 4 and fewer in groups 1 to 3 as compared to group 5 (see Table 5). The proportion of females classified in each of the five groups were 80%, 90%, 76%, 49%, and 68% respectively. The only other significant covariate was reading literacy, which was lower in group 2. When considering other groups as the reference (data not shown), it appeared that groups 1, 2 and 4 significantly differed in terms of gender, while groups 4 and 5 differed from group 2 regarding the reading literacy level, with a significantly lower reading literacy level in group 2. On the other hand, the residence, socio-economic and family wealth covariates were never significant, even though some coefficients were very close to significance, especially the coefficient of the critical life events covariate in the case of group 5.

Group 4 comprised the respondents with lower overall somatic complaint scores and with relatively low changes between periods, that is the individuals with the overall lowest level of somatic complaints. Both lags of the dependent variable and the critical life events covariate were significant at the visible level. Compared to group 4, group 5 comprised respondents with a slightly higher variability of scores over time, while average scores varied more and were significantly higher in a number of cases. On the other hand, groups 1 to 3 comprised respondents with substantially more complex trajectories of somatic complaints: both their variability and their overall level was higher, especially in group 1, and no influence of the

Table 5 Main characteristics of the respondents classified in the five groups of the final model

Variables	Categories	Groups				
		1	2	3	4	5
Gender	Female	80.4%	90.0%	75.5%	49.0%	67.8%
	Male	19.6%	10.0%	24.5%	51.0%	37.2%
Country of birth	Switzerland	87.0%	86.7%	92.6%	95.2%	95.8%
	Other	13.0%	13.3%	7.4%	4.8%	4.2%
Academic track attended at lower secondary education level	High	58.7%	40.0%	52.5%	45.6%	50.6%
	Extended	26.1%	30.0%	28.9%	36.0%	31.1%
	Basic	15.2%	30.0%	18.7%	18.4%	18.4%
Residence	Rural area	17.4%	33.3%	38.7%	38.0%	38.3%
	Urban area	82.6%	66.7%	61.3%	62.0%	61.7%
(PISA) Reading literacy	Min.: 0	3.35	3.10	3.46	3.53	3.53
	Max.: 5	(0.92)	(1.09)	(0.96)	(0.98)	(1.02)
Highest parental ISEI	Min.: 16	54.70	48.90	52.84	54.15	52.78
	Max.: 90	(15.71)	(16.67)	(15.53)	(15.70)	(15.15)
Family wealth	Min.: -2.93	0.20	-0.03	-0.01	0.14	0.00
	Max.: -3.38	(0.74)	(0.82)	(0.75)	(0.73)	(0.77)

Note: We display, separately for each group, the percentage of each category for categorical variables, and the mean and standard deviation for numerical variables.

critical life events covariate was observed. Moreover, the individuals classified into these three groups had generally one or several periods with a high level of somatic complaints. Both lags of the dependent variable were significant in group 2, while only the first lag was significant in groups 1 and 3, indicating that in these two latter groups, the past levels of somatic complaints had less effect on the current level.

Lastly, we compared the final clustering with the variables measuring the level of substance consumption (Table 6). Overall, the level of association was not very strong, but it was nevertheless highly significant in many cases. The association with alcohol consumption tended to change rapidly from one period to the next, but with no discernible trend. Similar findings were observed for cannabis consumption, where significant and non-significant associations alternated. The pattern of tobacco use was more distinct: the association with the five groups of the clustering was always significant except for 2004 and 2010. Compared to groups 4 and 5, daily smokers were represented at a substantially higher proportion in group 1, and to a lesser extent in groups 2 and 3 (data not shown). Finally, the association between the five groups and the use of tranquilizers and sleeping pills was highly significant for each period. Even if the consumption level was rarely higher than 1–3 times per month in all groups, and if most respondents did not consume at all, the average consumption level was significantly higher among respondents clustered into group 2 (data not shown).

Table 6 Relationships between the five groups of the final model and substance use

Substances		2001	2002	2003	2004	2005	2006	2007	2010	2014
Alcohol	V	0.050	0.060	0.080	0.050	0.080	0.070	0.070	0.080	0.090
	p	0.706	0.441	0.018	0.758	0.011	0.057	0.200	0.035	0.002
Tobacco	V	0.090	0.080	0.090	0.070	0.080	0.090	0.090	0.070	0.080
	p	0.001	0.039	0.002	0.107	0.016	0.001	0.004	0.194	0.010
Cannabis	V	0.060	0.080	0.080	0.080	0.070	0.060	0.090	0.060	0.080
	p	0.466	0.045	0.035	0.018	0.121	0.483	0.001	0.447	0.007
Tranquilizers and sleeping pills	V	0.130	0.100	0.160	0.100	0.130	0.100	0.160	0.130	0.090
	p	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.001

Note: For each wave of the study, we provide the Cramer's V measure giving the level of association between groups and substance use, and the corresponding p-value.

## 4 Discussion

The main finding of this study was the identification of several distinct groups of somatic complaints trajectories based on a scale representing the sum of eight different complaints. These trajectories remained distinct throughout the entire observation

period covered by the TREE data, which is from age 16 to age 30. As these trajectories already differed at age 16, we can hypothesize that factors already present during childhood, and thus beyond the control of youths themselves, may be the cause of such a differentiation. Since it is known that a higher level of somatic complaints is associated with subsequent health issues, we can conclude that 1) some groups of adolescents in Switzerland were experiencing a situation of vulnerability beginning before adolescence, and that 2) this condition is likely to persist even beyond the period covered by our study. A second conclusion is that if critical life events are related to somatic complaints, this relationship was visible only among individuals with low levels of somatic complaints. For their counterparts with high levels of somatic complaints, the impact of critical life events could have been masked by the inherent variability observed in somatic complaint scores. This brings us back to the fact that even if critical life events experienced during the transition period of adolescence and young adulthood can have an impact on somatic complaints, this impact remains limited, and the most important causes of a high level of somatic complaints are to be found elsewhere. This also leads us to the assumption that if somatic complaints are triggered by the occurrence of critical life events, the influence of such events is mainly of short duration, as the somatic complaint score often decreased one period later. Thirdly, the consumption of tobacco and tranquilizers and sleeping pills was significantly associated with the typology of somatic complaint trajectories: higher substance consumption was associated with the groups reporting the highest overall somatic scores. As most substance use began during the period covered by this study rather than before, it should not be considered as a cause of somatic complaints but rather as a consequence, especially in the case of tranquilizers and sleeping pills, which can be used to relieve some of these complaints.

Returning to the five hypotheses of the Introduction section, we can conclude that the four hypotheses regarding the role of gender, less than average academic achievement, critical life events, and substance use were at least partially confirmed, while hypothesis 2 related to the influence of a low SES context was not, at least at the multivariate level. This last result was not expected, but it may be related to the specific Swiss context: socio-economic differences may be weaker than in other countries, a situation reinforced by an overall high-level education system offering opportunities to all adolescents. Therefore, the socio-economic background of an individual could lose part of its importance.

The typology of somatic complaint trajectories identified in this study illustrates both the importance and the long lasting aspects of somatic complaints: important differences were observed between respondents classified into each group, with clearly differentiated overall somatic complaint levels. Moreover, even though both inter- and intra-subject variability may be high, many respondents classified in the first three groups stayed at a high level of reported complaints during the entire period of observation, i. e. from 16 to 30 years. In terms of life course, that

means that the presence of somatic complaints early in life is susceptible to deploy effects during the whole adolescence and (at least) the beginning of adulthood. Since many important determinants for the entire life (such as entry on the labour market, and the beginning of a steady relationship) are also taking ground during the same period, somatic complaints could be a very important indicator for the success or not of the entire life of an individual.

Significant associations were found at three different levels. First, the level of somatic complaint measured in a given year was always related to the level observed in the previous years. This is a clear indication that somatic complaints persist over time. Second, the introduction of fixed covariates representing the prevailing situation at the beginning of the observed period improved the clustering process. The level and development of somatic complaints over time were influenced not only by gender, but also by academic achievement operationalized here through the reading literacy level achieved at the end of mandatory school. While gender is a factor exogenous to the youths themselves, both exogenous and endogenous factors are at work in the case of reading literacy, as a large and solid body of research suggests that social origin substantially influences the academic achievement of children (Boudon 1973; Bourdieu 1977; Sacchi et al. 2011). It is also worth noting that even though females may tend to report a higher number of somatic complaints than males, our clustering shows that not only the average level, but also the variability from one period to the next was higher in the groups comprising more females. One could make the hypothesis that during the period of life covered by our data, there is more pressure on females than on males regarding the need to conciliate studies and work on one side, and family and children on the other. The level and variability of somatic complaints could then be a revealer of the difficulty of reconciling these two aspects of life.

Third, the covariate added at the visible level, i. e. the critical life events score, was also significant, at least for one of the five groups of the clustering. It may seem surprising to have found this association only in the group with the lowest variability and lowest average score (group 4). However, this finding can be interpreted in the light of the overall high variability of somatic complaint scores observed from one year to the next. Since the number of somatic complaints measured in this study was limited, youth reporting a high level of somatic complaints were not likely to increase this level further when a critical life event occurred. Putting it differently, the effect of a critical life event on somatic complaints, even if possible, is masked by the somatic complaints already present. On the other hand, when looking at young people who did not have many somatic complaints, the occurrence of a critical life event made a difference. This is corroborated by the fact that in addition to the significance of critical life events observed in group 4, this covariate was also very close to being significant for group 5, the second group with the lowest overall variability of somatic complaints.

Regarding the use of substances, tobacco and tranquilizers and sleeping pills were especially associated with the different trajectories of somatic complaints. Both kinds of substances were more prominently used in the first three groups of the clustering, those with the highest somatic complaint levels. It may seem plausible to observe a higher level of tranquilizer and sleeping pill consumption among groups presenting higher levels of somatic complaints. However, the relationship is more complex, given that tranquilizer and sleeping pill consumption was higher in group 2 than in groups 1 and 3. It can then be hypothesized that the exact nature of the somatic complaints experienced by the respondents in the different groups is the reason for the differentiated consumption of tranquilizers and sleeping pills.

Vulnerability is a dynamic multidimensional situation related to health, but also to other dimensions such as education, family, economic conditions and time (Spini et al. 2017). Exact causes of vulnerability are difficult to identify, as well as the precise moment, if any, of entry into a vulnerable situation. Nevertheless, it is well established that vulnerability is a long-lasting condition that can begin at any age. In this study, we explored the development of specific health issues, somatic complaints, which may be considered as conditions leading to or fostering the development of vulnerability. We were particularly interested in the fact that somatic complaints could appear very early in the life course, during childhood or adolescence. Therefore, they may affect individuals during their whole life, and influence the way an individual will succeed in terms of educational achievement, entering the labour market, founding a family, etc. Understanding the early development and trajectories of somatic complaints is therefore of great importance.

This analysis draws on longitudinal data covering a period of fifteen years, which allows us to analyse long term trajectories of somatic complaints. However, our analysis is subject to some limitations. A first limitation lies in the restriction to respondents with complete data sequences. Even though the resulting number of sequences ( $n = 1\,161$ ) was not a limitation for the statistical computations, there was an obvious selection bias, because individuals who had less linear trajectories, dropped out of school, had problems with legal authorities or belonged to a low socio-economic stratum were more likely to not respond to one or several survey waves of the TREE panel study, or to drop out of the study altogether. One solution to this problem would have been to impute at least part of missing data, as has been done in a similar analysis on cannabis and tobacco consumption trajectories (Berchtold and Surís 2017) but there is still much uncertainty about the best approach to adopt. Using data from a real survey, we compared different strategies combining multiple imputation and the chained equations method, the two main objectives being 1. However, when analysing individual trajectories, this approach is adequate only if trajectories are smooth, with little changes from one wave to another. This is not the case with somatic complaints. The use of multiple imputations could have reduced this issue, but at the cost of having to optimize several times a HMTD model with



the same structure on different datasets, an especially difficult task (Taushanov and Berchtold 2017). As our objective was not to establish the prevalence of somatic complaints and their different trajectories in the Swiss population, sample bias does not compromise our results. We can assume that in a non-biased sample more complex trajectories and higher levels of somatic complaints would be found. Another possible bias lies in the irregular intervals between survey waves (yearly intervals for the first seven waves, three to four year intervals for the subsequent waves). In our modelling, we did not take into account this data feature. However, previous HMTD modelling of data with similar characteristics proved that adjustment for unequally spaced data points did not really improve the results (Bolano 2015). Finally, we relied on a global indicator aggregating the number and importance of eight different somatic symptoms. However, previous research has shown that different subgroups of somatic complaints do exist, and that these subgroups can differ between men and women (Tsai 2010). Therefore, future studies should take into account the multidimensionality of somatic complaints in order to identify more specific trajectories. Moreover, the measurement tool used here to establish the level of somatic complaints may not be the most used, but many tools were developed in the past and the tool we used shares many characteristics with them: each complaint is evaluated on a Likert scale, and the most usual complaints such as headache or stomach ache are evaluated (Zijlema et al. 2013).

This study breaks the ground for further research on the development of somatic complaints throughout the life course and their association with personal and socio-economic factors. As it has been shown, significant differences in terms of somatic complaints can be observed as early as age 16, and these differences persist in the long term. Therefore, additional studies should endeavour to trace back the emergence of somatic complaints to childhood in order to identify the first and main causes of such a risk to develop a vulnerable condition.

## 5 References

- Barkmann, Claus, Christiane Otto, Gerhard Schön, Michael Schulte-Markwort, Robert Schlack, Ulrike Ravens-Sieberer, Fionna Klasen, and BELLA Study Group. 2015. Modelling Trajectories of Psychosomatic Health Complaints in Children and Adolescents: Results of the BELLA Study. *European Child & Adolescent Psychiatry* 24(6): 685–694.
- Barsky, Arthur J., Heli M. Peekna, and Jonathan F. Borus. 2001. Somatic Symptom Reporting in Women and Men. *Journal of General Internal Medicine* 16(4): 266–275.
- Bekhuis, Ella, Lynn Boschloo, Judith G.M. Rosmalen, Marrit K. de Boer, and Robert A. Schoevers. 2016. The Impact of Somatic Symptoms on the Course of Major Depressive Disorder. *Journal of Affective Disorders* 205: 112–118.
- Berchtold, André. 2002. High-Order Extensions of the Double Chain Markov Model. *Stochastic Models* 18(2): 193–227.

- Berchtold, André and Joan-Carles Surís. 2017. Imputation of Repeatedly-Observed Multinomial Variables in Longitudinal Surveys. *Communications in Statistics – Simulation and Computation* 46(4): 3267–3283.
- Bolano, Danilo. 2015. *Markovian Modelling of Life Course Data*. University of Geneva. <http://archive-ouverte.unige.ch/unige:74459>.
- Bolano, Danilo and André Berchtold. 2016. General Framework and Model Building in the Class of Hidden Mixture Transition Distribution Models. *Computational Statistics & Data Analysis* 93: 131–145.
- Boudon, Raymond. 1973. *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*. New York: Wiley.
- Bourdieu, Pierre. 1977. Cultural Reproduction and Social Reproduction. Pp. 487–511 in *Power and Ideology in Education*, edited by Pierre Bourdieu. New York: Oxford University Press.
- Choi, Laura. 2009. Financial Stress and Its Physical Effects on Individuals and Communities. *Community Development Investment Review* 3: 120–122.
- Compas, Bruce E., Barry M. Wagner, Lesley A. Slavin, and Kathryn Vannatta. 1986. A Prospective Study of Life Events, Social Support, and Psychological Symptomatology During the Transition From High School to College. *American Journal of Community Psychology* 14(3): 241–257.
- Dey, Michelle, Anthony F. Jorm, and Andrew J. Mackinnon. 2015. Cross-Sectional Time Trends in Psychological and Somatic Health Complaints Among Adolescents: A Structural Equation Modelling Analysis of “Health Behaviour in School-Aged Children” Data From Switzerland. *Social Psychiatry and Psychiatric Epidemiology* 50(8): 1189–1198.
- Eliassen, Marie, Svend Kreiner, Jeanette F. Ebstrup, Chalotte H. Poulsen, Cathrine J. Lau, Sine Skovbjerg, Per K. Fink, and Torben Jørgensen. 2016. Somatic Symptoms: Prevalence, Co-Occurrence and Associations With Self-Perceived Health and Limitations Due to Physical Health – A Danish Population-Based Study. *PLOS ONE* 11(3): e0150664.
- Ganzeboom, Harry B. G., Paul M. de Graaf, Donald J. Treiman, Jan de Leeuw, H. B. G. Ganzeboom, P. M. de Graaf, D. J. Treiman, and J. de Leeuw. 1992. *A Standard International Socio-Economic Index of Occupational Status*. Tilburg University, Work and Organization Research Centre. <http://EconPapers.repec.org/RePEc:tiu:tiuwor:85970031-d601-46e3-befb-156cd78a09d9> (02.08.2017).
- Goodwin, Guy M. 2006. Depression and Associated Physical Diseases and Symptoms. *Dialogues in Clinical Neuroscience* 8(2): 259–265.
- Grob, Alexander, Ruth Lüthi, Florian G. Kaiser, August Flammer, Andrew Mackinnon, and Alex J. Wearing. 1991. Berner Fragebogen zum Wohlbefinden Jugendlicher (BFW). *Diagnostica* 37(1): 66–75.
- Härmä, Ann-Mari, Riittakerttu Kaltiala-Heino, Matti Rimpelä, and Päivi Rantanen. 2002. Are Adolescents With Frequent Pain Symptoms More Depressed? *Scandinavian Journal of Primary Health Care* 20(2): 92–96.
- Hart, Shayla L., Stacy C. Hodgkinson, Harolyn M. E. Belcher, Corine Hyman, and Michele Cooley-Strickland. 2013. Somatic Symptoms, Peer and School Stress, and Family and Community Violence Exposure Among Urban Elementary School Children. *Journal of Behavioral Medicine* 36(5): 454–465.
- Jonsson, Frida, Anne Hammarström, and Per E. Gustafsson. 2014. Social Capital Across the Life Course and Functional Somatic Symptoms in Mid-Adulthood. *Scandinavian Journal of Public Health*, September, 42(7): 581–588.
- Kapfhammer, Hans-Peter. 2006. Somatic Symptoms in Depression. *Dialogues in Clinical Neuroscience* 8(2): 227–239.
- Mohapatra, Satyakam, Sardar J. K. Deo, Ashirbad Satapathy, and Neelmadhav Rath. 2014. Somatoform Disorders in Children and Adolescents. *German Journal of Psychiatry* 17(1): 19–24.

- Nakao, Mutsuhiro and Eiji Yano. 2006. Somatic Symptoms for Predicting Depression: One-Year Follow-up Study in Annual Health Examinations. *Psychiatry and Clinical Neurosciences* 60(2): 219–225.
- Poikolainen, Kari, Riitta Kanerva, and Jouko Lönnqvist. 1995. Life Events and Other Risk Factors for Somatic Symptoms in Adolescence. *Pediatrics* 96(1): 59–63.
- Radke-Yarrow, Marian and Earnestine Brown. 1993. Resilience and Vulnerability in Children of Multiple-Risk Families. *Development and Psychopathology* 5(4): 581–592.
- Raftery, Adrian. 1995. Bayesian Model Selection in Social Research (with Discussion by Andrew Gelman & Donald B. Rubin, and Robert M. Hauser, and a Rejoinder). *Sociological Methodology* 25: 111–163.
- Sacchi, Stefan, Sandra Hupka Brunner, Barbara Stalder, and Markus Gangl. 2011. The Impact of Social Origin and Migration Background on Transition Into Post Compulsory Education and Training. Pp. 120–156 in *Youth Transitions in Switzerland: Results From the TREE Panel Study*, edited by Manfred Max Bergman, Sandra Hupka-Brunner, Anita Keller, Thomas Meyer, and Barbara Elisabeth Stalder. Zurich: Seismo.
- Sek-yum, Steven Ngai, Lam Ching Man, and Choy Bing Kong. 1992. Stressful Life Events, Perceived Stress and Somatic Problems Among Young People in Hong Kong. *Asia Pacific Journal of Social Work and Development* 2(1): 27–44.
- Sigmon, Sandra T., Jennifer J. Pells, Nina E. Boulard, Stacy Whitcomb-Smith, Teresa M. Edenfield, Barbara A. Hermann, Stephanie M. LaMattina, Janell G. Scharrel, and Elizabeth Kubik. 2005. Gender Differences in Self-Reports of Depression: The Response Bias Hypothesis Revisited. *Sex Roles* 53(5–6): 401–411.
- Spini, Dario, Laura Bernardi, and Michel Oris. 2017. Toward a Life Course Framework for Studying Vulnerability. *Research in Human Development* 14(1): 5–25.
- Surís, Joan-Carles, Christina Akre, Claire Piguet, Anne-Emmanuelle Ambresin, Grégoire Zimmermann, and André Berchtold. 2014. Is Internet Use Unhealthy? A Cross-Sectional Study of Adolescent Internet Overuse. *Swiss Medical Weekly* 144: w14061.
- Taushanov, Zhivko and André Berchtold. 2017. A Direct Local Search Method and Its Application to a Markovian Model. *Statistics, Optimization & Information Computing* 5: 19–34.
- TREE. 2016. *Documentation on the First TREE Cohort (TREE1), 2000–2016*. Bern: TREE. [http://www.tree.unibe.ch/unibe/portal/fak\\_wiso/c\\_dep\\_sowi/micro\\_tree/content/e206328/e305140/e305154/files476810/TREE\\_2016\\_Project\\_documentation\\_TREE1\\_2000-2016\\_English\\_ger.pdf](http://www.tree.unibe.ch/unibe/portal/fak_wiso/c_dep_sowi/micro_tree/content/e206328/e305140/e305154/files476810/TREE_2016_Project_documentation_TREE1_2000-2016_English_ger.pdf) (09.02.2017).
- Tsai, Chung-Huang. 2010. Factor Analysis of the Clustering of Common Somatic Symptoms: A Preliminary Study. *BMC Health Services Research* 10: 160.
- Vassend, Olav. 1989. Dimensions of Negative Affectivity, Self-Reported Somatic Symptoms, and Health-Related Behaviors. *Social Science & Medicine* 28(1): 29–36.
- Verhoof, Eefje, Heleen Maurice-Stam, Hugo Heymans, and Martha Grootenhuys. 2012. Growing Into Disability Benefits? Psychosocial Course of Life of Young Adults With a Chronic Somatic Disease or Disability. *Acta Paediatrica* 101(1): e19–26.
- Wright, Majjaliisa Rauste-von and Johan von Wright. 1981. A Longitudinal Study of Psychosomatic Symptoms in Healthy 11–18 Year Old Girls and Boys. *Journal of Psychosomatic Research* 25(6): 525–534.
- Yoshimasu, Kouichi. 2012. Substance-Related Disorders and Somatic Symptoms: How Should Clinicians Understand the Associations? *Current Drug Abuse Reviews* 5(4): 291–303.
- Zijlema, Wilma L., Ronald P. Stolk, Bernd Löwe, Winfried Rief, BioSHaRE, Peter D. White, and Judith G. M. Rosmalen. 2013. How to Assess Common Somatic Symptoms in Large-Scale Studies: A Systematic Review of Questionnaires. *Journal of Psychosomatic Research* 74(6): 459–468.